Sungani, H., Ngatunga, B. P., Koblmüller, S., Mäkinen, T., Skelton, P. H., & Genner, M. J. (2017). Multiple colonisations of the Lake Malawi catchment by the genus Opsaridium (Teleostei: Cyprinidae). *Molecular Phylogenetics and Evolution*, *107*, 256-265. https://doi.org/10.1016/j.ympev.2016.09.027

Multiple colonisations of the Lake Malawi catchment by the genus *Opsaridium* (Teleostei: Cyprinidae)

Harold Sungani [a, b,*], Benjamin P. Ngatunga [c], Stephan Koblmüller [d], Tuuli Mäkinen [e], Paul H. Skelton [e] and Martin J. Genner [a]

[a] School of Biological Sciences, Life Sciences Building, 24 Tyndall Avenue, Bristol. BS8 1TQ. United Kingdom.

[b] Fisheries Research Unit, PO Box 27, Monkey Bay, Malawi.

[c] Tanzania Fisheries Research Institute, PO. Box 9750, Dar es Salaam, Tanzania.

[d] Institute of Zoology, University of Graz, Universitätsplatz 2, A-8010 Graz, Austria.

[e] South African Institute for Aquatic Biodiversity, Private Bag 1015, Grahamstown 6140, South Africa.

*Corresponding author at. School of Biological Sciences, Life Sciences Building, 24 Tyndall Avenue, Bristol. BS8 1TQ. United Kingdom.

Email address: harold.sungani@bristol.ac.uk (H. Sungani).

Abstract

It has been proposed that the fish faunas of African rivers assemble through multiple colonisation events, while lake faunas form additionally through intralacustine speciation. While this pattern has been established for many lineages, most notably cichlids, there are opportunities to further investigate the concept using phylogenies of congeneric endemic species within ancient lake catchments. The Lake Malawi catchment contains three river-spawning cyprinids of the genus *Opsaridium*, two of which are endemic. These species differ in body size, migratory behaviour and habitat use, but it has never previously been tested if these represent a monophyletic radiation, or have instead colonised the lake independently. We placed these species in a broader phylogeny of *Opsaridium* and the related genus *Raiamas*, including all known species from the river systems surrounding Lake Malawi. Our results suggest that each of the species has independently colonised the lake catchment, with all three taxa having well-defined sister taxa outside of the lake, and all sharing a common ancestor ~14.9 million years ago, before the Lake Malawi basin started to form ~8.6 million years ago. Additionally, the results strongly support previous observations that *Opsaridium* is not a monophyletic group, but instead contains *Raiamas* from the Congo drainage. Together these results are supportive of the concept that river fish faunas within African catchments are primarily assembled through a process of accumulation from independent origins, rather than within-catchment speciation and adaptive radiation. In light of these results we also suggest there is scope for a re-evaluation of systematics of both *Opsaridium* and *Raiamas*.

Key words. Molecular phylogeny, Chedrina, phylogeography, river fishes, East Africa

1. Introduction

It is estimated that the African freshwater fish fauna contains approximately 3000 species (Lévêque et al., 2008). A substantial component of this diversity can be found in the adaptive radiations of the East African Rift Valley lakes, while neighbouring river systems typically have lower diversity over equivalent spatial scales. Given that rivers and large lakes can both have a broad range of ecological niches, it is possible that the distribution of species diversity is not exclusively governed by niche availability, and instead the contrasting patterns of species diversity may have arisen as a result of a different community assembly processes operating within these different habitats (Seehausen, 2015). Lakes tend to be relatively stable environments over evolutionary timescales, potentially allowing high species diversity to arise through intralacustrine speciation and adaptive niche evolution. By contrast riverine environments tend to be more dynamic and unstable over evolutionary timescales, potentially giving a relatively greater role to immigration and between-catchment allopatric divergence (Seehausen, 2015).

Evidence in support of different assembly processes operating in lakes and rivers is most prominent in African cichlid fishes. Cichlids dominate the faunas of the East African lakes, and there is extensive evidence of intralacustrine diversification in most of the lake-restricted species groups (Seehausen, 1996; Snoeks, 2004; Koblmüller et al., 2008a; Salzburger et al., 2014). By contrast riverine cichlid assemblages in systems surrounding the Great Lakes of East Africa often are comprised of only a handful of geographically widespread species, albeit with some phenotypic divergence in allopatry (Katongo et al., 2007; Koblmüller et al., 2012; Banyankimbona et al., 2013; Egger et al., 2015; Meyer et al., 2015; Nichols et al., 2015). In some cases this has required description of distinct species, for example *Orthochromis* from the Lake Tanganyika drainage (De Vos and Seegers 1998). Further evidence that speciation tends to be constrained in rivers relative to lakes is present in the weakly radiating non-cichlid lineages from Lake Malawi [clariid catfishes (Agnèse and Teugels, 2001)] and Lake Tanganyika [*Synodontis* catfishes (Day et al., 2013; Pinton et al., 2013); claroteine catfishes (Peart et al., 2014); mastacembelid eels (Brown et al., 2010)]. Most African fish families, however, have failed to radiate in lakes to any degree and their diversity tends to be greatest in riverine or still-water pool habitats. In these lineages sister species often have allopatric distributions [e.g. Dorn et al. (2011) for *Nothobranchius* killifishes, Schmidt et al. (2014) for

Chiloglanis catfishes], implying a strong role for geographical separation in the speciation of non-lacustrine freshwater fish lineages across the continent.

Lake Malawi (=Nyasa) is estimated to have formed between ~4.5 and 8 Ma (Danley et al., 2012). The catchment contains at least 450 species of cichlid fish of which only five are frequently found in multiple river systems (Snoeks, 2004). In addition to this diversity ~50 other non-cichlid fish species are known from the catchment, of which most are non-endemics (Snoeks, 2004). One of the less well-studied species groups, from a phylogenetic perspective, is that of three cyprinids in the genus *Opsaridium*. These species differ extensively in body size, migratory behaviour and habitat use. Two are large-bodied species endemic to the catchment and migrate from the lake to rivers to breed [mpasa, *Opsaridium microlepis* (Günther, 1864); sanjika, *Opsaridium microcephalum* (Günther, 1864)]. A third non-endemic species, dwarf sanjika, *Opsaridium tweddleorum* Skelton 1996, is small-bodied in comparison, and only found in streams and rivers. It has never previously been tested if these represent a monophyletic *in-situ* radiation, or have instead colonised the lake independently.

Opsaridium comprises part of the "chedrin" group. This group has been defined as the Chedrini, one of three tribes within the subfamily Danioninae alongside Danionini and Rasborini (Tang et al., 2010; Ahnelt et al., 2015). It has also been defined as the Chedrina, a subtribe of the tribe Danionini, within the subfamily Danioninae (Liao et al., 2011). The chedrin group has a broad distribution across Africa and Asia, and a combination of molecular and morphological evidence (Tang et al., 2010; Liao et al., 2011, Liao et al., 2012) suggests it comprises 17 genera. Eight of these genera are from Africa (*Chelaethiops*, *Engraulicypris*, *Leptocypris*, *Mesobola*, *Neobola*, *Opsaridium*, *Raiamas* and *Rastrineobola*), while 10 are from Asia (*Barilius*, *Bengala*, *Cabdio*, *Luciosoma*, *Malayochela*, *Nematabramis*, *Opsarius*, *Raiamas*, *Salmostoma* and *Securicula*; generic names following Eschmeyer et al., 2016). Thus, only one genus, *Raiamas*, is found in both continents, but notably no species are shared between Africa and Asia. Morphological evidence has resolved monophyly of eight genera (*Nematobramis*, *Luciostoma*, *Bengala*, *Opsarius*, *Raiamas*, *Salmostoma*, *Neobola* and *Chelaethiops*) with respect to other representatives of the Danioninae (Liao et al., 2011). Molecular phylogenetic evidence has also repeatedly resolved monophyly of these 17 chedrin genera (Rüber et al., 2007; He et al., 2008; Tang et al., 2010; Liao et al., 2012), with the exception of the placement of *Raiamas guttatus* in a study by Fang et al. (2009). The chedrin group has also been proposed to include *Esomus* based on morphological evidence (Liao et al.,

2011, Liao et al., 2012), but the placement is not supported by molecular phylogenetic evidence (Tang et al., 2010).

Notably, molecular phylogenies have resolved African chedrins as monophyletic with respect to the Asiatic chedrins, demonstrating that *Raiamas* is a non-monophyletic group (Tang et al., 2010). Additionally, there have been indications that *Opsaridium* is non-monophyletic (Tang et al., 2010). Here we place the three *Opsaridium* from the Lake Malawi catchment into a time-calibrated phylogeny of African and Asiatic chedrins, including *Opsaridium* from neighbouring catchments. Our results provide further evidence that river faunas assemble through independent immigration events, rather than within-catchment radiation. The results are considered in light of the inferred age of Lake Malawi, and the current taxonomy of *Opsaridium* and *Raiamas*.

2. Methods

2.1 Taxonomic sampling

Our goal was to reconstruct a phylogeny of Lake Malawi catchment *Opsaridium* species within the broader context of the African and Asiatic chedrin species. Thus, we included data for nine of the 12 valid species in *Opsaridium* (Table 1). We used published data for four species, *Opsaridium boweni* (Fowler, 1930), *Opsaridium ubangiense* (Pellegrin, 1901), *Opsaridium zambezense* (Peters, 1852) and *Opsaridium peringueyi* (Gilchrist and Thompson, 1913). We generated new data for a further five species, *Opsaridium leleupi* (Matthes, 1965), *Opsaridium loveridgii* (Norman, 1922), *Opsaridium microlepis, Opsaridium microcephalum* and *Opsaridium tweddleorum*. We also generated new sequences for *Opsaridium zambezense*. Samples for these six species were collected from scientific surveys and artisanal fisheries. Three species were not available for study, namely *Opsaridium engrauloides* (Nichols, 1923), *Opsaridium maculicauda* (Pellegrin, 1926) and *Opsaridium splendens* Taverne and De Vos, 1997. Given evidence that some *Raiamas* are closely related to *Opsaridium* (Liao et al., 2012), we included data from nine of the 17 valid species, including the two Asiatic species *Raiamas guttatus* (Day 1870) and *Raiamas bola* (Hamilton, 1822), and seven African species *Raiamas moorii* (Boulenger 1900), *R. salmolucius* (Nichols and Griscom 1917), *R. christyi* (Boulenger 1920), *R. kheeli* Stiassny, Schelly and Schliewen 2006, *R. buchholzi* (Peter 1876), *R. batesii* (Boulenger 1914) and *R. senegalensis* (Steindachner 1870). Eight species were not available

for study, namely *R. ansorgii* (Boulenger 1910), *R. intermedius* (Boulenger 1915), *R. levequei* Howes and Teugels 1989, *R. longirostris* (Boulenger 1902), *R. nigeriensis* (Daget 1959), *R. scarciensis* Howes and Teugels 1989, *R. shariensis* (Fowler 1949) and *R. steindachneri* (Pellegrin 1908). To place these *Opsaridium* and *Raiamas* data in a broader perspective, we included representatives of other chedrin genera from Africa (*Engraulicypris*, *Mesobola, Leptocypris, Neobola*, *Chelaethiops*), and Asia (*Opsarius, Barilius, Salmostoma, Securicula, Cabdio, Lucisoma, Malayochela* and *Nematobramis*). Two more distantly related cyprinids were employed as outgroups (*Danio* and *Barbus*). Finally, we included two catostomid genera (*Hypentelium* and *Ictiobus*), due to the close relationship of the Cyprinidae and Catostomidae (Near et al., 2012). This enabled the use of a fossil calibration of the earliest representative of the Catostomidae subfamily Ictiobinae to date the divergence between the Ictiobinae and the Catostominae, following Near et al. (2012).

2.2 *DNA extraction and sequencing*

Total genomic DNA was extracted from tissues (muscle and fin clips preserved in 95% EtOH and stored frozen) using the Wizard genomic DNA purification 106 kit 119 (Promega), following the protocol of the manufacturer. We amplified the nuclear recombination-activating gene (*RAG1*) gene using the forward primer 2598F (5'-CCA ACC CCT GCA CAC TCT ACG T-3') and reverse primer 4067R (5'-TCA AAC GTT TTG GAC TGC CTT GCA TT-3') from Liao et al. (2012). PCR consisted of an initial denaturation of 95°C for 5 min for RAG1, followed by 35 cycles of 95°C for 1 min, annealing at 58°C for 1 min, and extension at 72°C for 2 min, before a final extension step of 72°C for 7 min. We amplified the mitochondrial cytochrome oxidase subunit 1 (*COI*) gene using the forward primer LCO1490 (5'-GGT CAA CAA ATC ATA AAG ATA TTG G-3') and reverse primer HCO2198 (5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3') from Folmer et al. (1994). PCR consisted of an initial denaturation of 95°C for 4 min, followed by 35 cycles of 95°C for 30s, annealing at 58°C for 30s, and extension at 72°C for 30s for *COI*, before a final extension step of 72°C for 7 min. Novel *RAG1* and *COI* sequences are listed in Table 1. Purified PCR products were sequenced on ABI3730XL sequencers at Macrogen (Amsterdam, NL). We also used previously published cytochrome b (*cytb*) and rhodopsin (*RH1*) sequences in our phylogenetic analyses (Table 2).

2.3 Phylogenetic analyses

Sequences for each gene were aligned using ClustalW in DAMBE 5.3.32 (Xia et al., 2013). Individual alignments for *COI*, *cytb*, *RAG1* and *RH1* consisted of 511, 826, 1308 and 704 nucleotides respectively. PartitionFinder (Lanfear et al., 2012) was used to identify the most likely data partitions for subsequent maximum likelihood (ML) and Bayesian inference (BI) analyses. ML analyses of the concatenated dataset were conducted with RAxML 8.2.0 (Stamatakis et al., 2008) with 100 bootstrap replicates, while BI analysis the concatenated dataset was conducted using a Markov Chain Monte Carlo (MCMC) algorithm in BEAST 1.8.2 (Drummond et al., 2012), with a lognormal relaxed clock model, and a run of 30 million generations recording every 10000 trees. The first 50% of these trees were removed as burn-in. Tracer 1.6 (Rambaut et al., 2013) was used to assess chain convergence to stationarity. Additionally we conducted BI analyses of the *COI*, *cytb*, *RAG1* and *RH1* genes independently, and the *COI* + *RAG1* together (Figs. S1-S5).

For divergence dating in the framework of the BI analysis in BEAST we employed two calibrations from a comprehensive molecular phylogeny of ray-finned fishes by Near et al. (2012), informed by evidence that the family Catostomidae is closely related to the family Cyprinidae. First we calibrated divergence of the catostomid subfamily Ictiobinae (represented by *Ictobus*) from the subfamily Catostominae (represented by *Hypentalium*). Following Near et al. (2012) we used the evidence of the first occurence of stem lineage Ictiobinae dated to 49.4 Ma, employing a lognormal prior with a mean of 0.764 and a standard deviation of 0.8 with 49.4 Ma as the minimum age offset. We set 57.0 Ma as a soft upper bound, based on fossil crown lineage Cypriniformes (see evidence in Near et al. 2012). Second, using the results of Near et al. (2012) we constrained the time of divergence between Cyprinidae and Catostomidae to a normal distribution with a mean divergence time of 98 Ma with a standard deviation of 11 Ma, giving a 95% probability range from 80 to 116 Ma, closely matching the range determined by Near et al. (2012). Note that the derived estimates should be considered approximations, as no internal calibrations are available and the application of calibrations points only at or close to the root is known to pull recent splits towards more ancient age estimates (Renner and Zhang, 2004)

The phylogenetic hypotheses of monophyly of the Lake Malawi *Opsaridium*, and monophyly of the genus *Opsaridium*, were tested using Shimodaira-Hasegawa (SH) tests. Maximum likelihood phylogenies with phylogenetic constraints imposed were constructed using RaxML

8.2.0. Site-wise log likelihoods for these trees and the maximum likelihood tree without constraints were then generated in RaxML. SH tests were then conducted in Consel (Shimodaira and Hasegawa, 2001).

2.4 Biogeography and ancestral area reconstruction

To reconstruct ancestral areas and the likelihood of dispersal and vicariance events we used Bayes-Lagrange (Statistical dispersal, extinction-cladogenesis, S-DEC) in RASP 3.2 (Yu et al., 2015), which account for topology, branch lengths and phylogenetic uncertainty. It was selected over other methods as it allowed us to assume dispersal events were only possible between adjacent catchments. We pruned the time-calibrated phylogeny (Fig. 1) to include one representative of each of the species in the "Africa clade". We then assigned distributions to each of the species using available distribution information (http://www.gbif.org/; Fig. 2). *Opsaridium zambezense* records listed in GBIF (http://www.gbif.org/) from south of the Zambezi, Pungwe and Buzi rivers were not included as these likely represent *O. peringueyi* (Skelton, 1996; Skelton, 2001).

Eight biogeographic areas were determined as: Lake Malawi catchment, Zambezi, Congo, Nile, East Africa, South-East Africa, West Africa and Lower Guinea (Fig. 3). These closely follow the African ichthyofaunal regions, delimited by Roberts (1975), which in turn are strongly influenced by catchment boundaries. Lake Malawi has a single outflow, into the Zambezi, but its fauna was treated as distinct to i) allow an investigation of whether the catchment has been colonised multiple times, and ii) due to the separation of the two systems by a series of rapids and waterfalls that has resulted in two distinct faunas (Tweddle et al., 1979, Tweddle & Willoughby, 1979). The Lake Malawi catchment is adjacent, but not directly connected to, the Congo and East African biogeographic areas, and is geographically distant from the other four areas. The Congo area also included the Lake Rukwa catchment, given evidence of a shared fauna over the last 25,000 years (Cohen et al., 2013). The South East Africa area (Limpopo and Save rivers southwards) was delimited due to apparent faunal separation from the more northerly Zambezi system (including the Buzi and Pungwe rivers). The East Africa area contains mainly rivers flowing east into the Indian Ocean. The Lower Guinea comprised multiple rivers flowing west into the Atlantic Ocean. The Nile system (including Lake Turkana) was treated as distinct from the West Africa region, which included Niger, Upper Guinea rivers and Lake Chad systems.

245

246 For the Bayes-Lagrange S-DEC analysis we used 1000 trees randomly sampled from the
247 posterior sample of trees obtained from the BI analysis in BEAST. We assumed an ancestral
248 range to include no more than four areas, given the maximum observed range of species is
249 three areas. We gave no prior range constraints, but we assumed dispersal constraints that
250 allowed only movement between adjacent catchments.

251

252 **3. Results**

253

254 *3.1 Phylogenetic relationships*

255

256 The ML phylogeny of the concatenated dataset was topologically similar to the BI consensus
257 tree of the same data (Fig. 1). African chedrins were resolved as a single monophyletic group
258 with an origin 19.97 Ma (95% CI range 14.46 to 26.21 Ma). Our focal *Opsaridium* genus was
259 resolved as paraphyletic with respect to *Raiamas salmolucius* and *Raiamas moori*. This broader
260 "*Opsaridium*" clade containing *R. salmolucius* and *R. moori* was dated to 14.85 Ma (95% CI
261 range 10.35 to 19.26 Ma).

262

263 Within the "*Opsaridium*" clade, all species with multiple specimens were resolved as
264 monophyletic. The Lake Malawi species did not comprise a monophyletic group. *Opsaridium*
265 *microcephalum,* an endemic of the Lake Malawi catchment, was resolved as a sister taxon to
266 *O. peringueyi* found in the eastward flowing rivers of southern Africa south of the Zambezi,
267 with an estimated divergence time of 2.87 Ma (95% CI range 1.46 to 4.63 Ma). *Opsaridium*
268 *microlepis*, another endemic of the Lake Malawi catchment, was resolved as a sister taxon of
269 *O. loveridgii* from the Rufiji-Ruaha basin, with an estimated divergence time of 10.24 Ma (95%
270 CI 5.91 to 14.97 Ma). *O. tweddleorum*, present only in Lake Malawi and the neighbouring
271 Lower Shire tributary of the Zambezi, was resolved as a sister taxon to *O. leleupi* from the SE
272 Congo, with an estimated divergence time of 8.65 Ma (95% CI range 4.56 to 13.82 Ma). The
273 phylogenetic hypotheses of monophyly of the Lake Malawi *Opsaridium,* and monophyly of
274 the genus *Opsaridium*, were both rejected (SH tests, both $P < 0.001$).

275

276 *3.2 Ancestral area reconstructions*

277

The results provided a strong indication that three *Opsaridium* present in the Lake Malawi catchment (*O. microcephalum*, *O. microlepis* and *O. tweddleorum*) have colonised the Lake Malawi biogeographic area independently. The most likely scenarios (66.0% probability; Fig 3) were that the common ancestor of *O. microcephalum*, *O. microlepis* and *O. tweddleorum* had a Congo distribution (66.0% probability; Fig 3), or a Malawi-Congo distribution (27.1% probability; Fig 3).

4. Discussion

4.1 Phylogenetic relationships of Opsaridium and Raiamas

Chedrins are a distinct component of the diverse cyprinid Danioninae subfamily. Within this group, previous work has strongly indicated the monophyly of the African clade, but paraphyly of the genus *Raiamas* with respect other African chedrins (Tang et al., 2010; Liao et al., 2012), and our results were consistent with this analysis. Previous analyses had placed *Opsaridium* as monophyletic [maximum likelihood and Bayesian inference in Liao et al. (2012)], or non-monophyletic with respect to *Chelaethiops bibie* (Tang et al., 2010), and *Raiamas salmolucius* [maximum parsimony in Liao et al. (2012)]. Our analyses have demonstrated clearer evidence of paraphyly of *Opsaridium* with respect to *R. salmolucius* and *R. moorii*.

Howes (1980) proposed twelve morphological characters to diagnose *Opsaridium*. Of these only four were confirmed by Stiassny et al. (2006), including an extended anal fin in mature males, well-developed axial lobes, large and granular tubercles, and a distinctively marked dorsal fin. A fifth character, a large dorsal fin with 13-15 branched rays was not mentioned by Stiassny et al. (2006). Further studies of more species of *Opsaridium* and *Raiamas* by Liao et al. (2012) demonstrated that none of these characters are diagnostic for *Opsaridium*. Our results, demonstrating paraphyly of *Opsaridium* with the inclusion of two *Raiama*s species, thus add weight to conclusions of Liao et al. (2012) that a comprehensive review of the African chedrin taxonomy is required.

4.2. *Origins of the Lake Malawi chedrins*

The Lake Malawi mpasa (*O. microlepis*) was found to share a most recent common ancestor with *O. loveridgii* from the neighbouring Ruaha-Rufiji system in Tanzania. It was estimated

312 that the divergence between these species took place on average 10.24 million years ago. The
313 two catchments are currently separated by the Kipengere/Livingstone mountain range that
314 originated as part of the western rift during the Late Miocene (15-30 million years ago; Burke
315 and Gunnell, 2008), with rifting continuing through the Pliocene. Evidence of a shared fish
316 fauna between the systems comes from the fossils of the Chiwondo beds in northern Malawi,
317 dated to between 2 and 3.75 Ma (Stewart and Murray, 2013). There, remains of claroteid
318 catfishes and tigerfish (*Hydrocynus*) are present, which are currently absent from the extant
319 Lake Malawi catchment fauna, and only occur together in one neighbouring drainage, the
320 Ruaha (Stewart and Murray, 2013). Molecular evidence from cichlids is also supportive of a
321 connection between the Malawi and Ruaha catchments between 2 to 7 million years ago
322 (Genner et al., 2015). It is notable that the area occupied by the ancestral species of *O.*
323 *microlepis* and *O. loveridgii* was ambiguous, with reconstructed areas including the Congo,
324 Malawi basin and East Coast all possible. On the basis of the combined evidence, it is possible
325 that the common ancestor of *O. microlepis* and *O. loveridgii* occupied a broader area across
326 central and eastern Africa.
327
328 The Lake Malawi catchment sanjika (*O. microcephalum*) was found to share a most recent
329 common ancestor with *O. peringueyi*, a species restricted to river systems south of the Zambezi.
330 The divergence was estimated to have taken place on average 2.87 million years ago. The
331 reconstructions of the range of common ancestor were ambiguous, with suggested ranges
332 encompassing the Congo, Malawi, Zambezi and South East coast areas. The apparent absence
333 of representatives of this clade from the geographically intermediate Zambezi system is
334 intriguing, particularly given the presence of *O. zambezense* in this system, suggesting suitable
335 habitat for *Opsaridium* is currently present. It is possible that competition with *O. zambezense*
336 contributed to the extirpation of the *O. microcephalum* - *O. peringueyi* lineage from the
337 Zambezi. Alternatively, or additionally, it is possible that historic events, such as the East
338 African megadroughts (Cohen et al., 2007; Moore and Eckardt, 2012), have extirpated
339 populations of the species. In this case it is possible that Lake Malawi may have acted as a
340 refuge for the *O. microcephalum* lineage during drought periods, particularly given evidence
341 the species is capable of spawning in fully lacustrine conditions (Tweddle & Turner 2014).
342
343 Dwarf sanjika (*O. tweddleorum*) is present in both Lake Malawi catchment and tributaries of
344 the Lower Shire River that is part of the Zambezi catchment. The species was found to share
345 common ancestry with *Opsaridium leleupi*, known only from the Upper Lualaba river (Fig 2),

346 with an estimated divergence time of 8.65 million years ago. Again, the ancestral area
347 reconstructions of this species were ambiguous, with the Congo, Zambezi and Lake Malawi
348 areas variously supported. It is likely the ancestral species had a broader distribution, but also
349 the apparent absence of representatives of the clade in geographically intermediate Zambezi
350 system is notable. This is again suggestive of historic extirpation of common ancestors from
351 Zambezi system, and that Lake Malawi may have acted as refuge during drought periods, in
352 turn allowing allopatric diversification.
353
354 It seems likely that the relationships within the "*Opsaridium*" clade will be further clarified
355 through sampling species not studied here, all of which have strongly restricted distributions
356 within tributaries of the Congo system. *Opsaridium engrauloides* is a species known only from
357 one type specimen of the Ubangui River at Bangui, Central African Republic. *Opsaridium*
358 *maculicauda* is known only from type specimens collected at Tshikapa in the Upper Kasai
359 system, Democratic Republic of the Congo. *Opsaridium splendens* is distributed in the
360 Malagarasi system of Tanzania and Burundi (Fig. 2). These species could be particularly
361 important given evidence that other fish species may have colonised the Malawi catchment
362 from the Congo system, possibly via the Luangwa river (Tweddle and Skelton, 2008; Egger et
363 al., 2015).
364
365 4.3. *Origins of the "Opsaridium" clade.*
366
367 Our analyses of ancestral areas revealed the Congo Basin as the most likely origin of much of
368 the "*Opsaridium*" clade (Fig. 3). The Congo is the most ancient drainage basin on the continent
369 with an age of at least 65 million years, although the present geographic structure of the basin
370 dates to ~5 million years ago (Stankiewicz and de Wit, 2006). The substantial catchment (ca.
371 3.8 mil km) will have ensured flow in downstream habitats even during regional droughts (e.g.
372 Dalibard et al., 2014). A recent study estimated that the basin harbours more than 1,250 species
373 of freshwater fish (Snoeks et al., 2011), so this stability may have helped to supporting high
374 phylogenetic diversity and species richness of riverine species present. However, whether the
375 basin has acted as a source of diversity for other regions of Africa is more equivocal, as
376 although several studies have considered historical phylogeography of African freshwater
377 fishes across broad spatial scales (e.g. Koblmüller et al., 2008b; Goodier et al., 2011; Brown et
378 al., 2010), few have reconstructed ancestral areas. Among the exceptions are studies of
379 *Synodontis* catfishes that have attributed basal nodes of the genus primarily to the West Africa,

the Nilo-Sudan or Congo regions, whereas the East African and Southern African taxa appear to be more derived (Day et al., 2013; Pinton et al., 2013). Notably, however, there was also evidence of Congo drainage species being derived from East and Southern African clades (Day et al., 2013; Pinton et al., 2013).

4.4. Assembly of the Lake Malawi catchment fauna

Over large spatial scales species richness within river systems is related to key environmental variables, including the catchment area, the energy input into the system, and the historical stability of the catchment (Oberdorff et al., 2011), including the frequency of river capture events. These factors in turn will affect the likelihood of the different processes that affect community assembly, most importantly immigration, speciation and extinction. Our study supports the hypothesis that river assemblages within the Lake Malawi catchment have formed through a process of immigration and within-catchment adaptation, rather than within-catchment speciation. This reinforces the importance of natural barriers within and across river networks in preventing gene flow, and allowing allopatric adaptive divergence to take place. The data are also supportive of periodic between-catchment dispersal events across African watersheds, which are often characterised by flat wetlands that periodically flood and allow fish movement between drainage systems (Beadle, 1974). The underlying reasons why most of the riverine fish lineages in the Lake Malawi catchment have failed to undergo speciation and adaptive radiation within these river systems remains unclear. One possibility is that inherent environmental instability in shallow river systems that surround large lakes favours generalist strategies and high dispersal that prevent the local adaptation and assortative mating that often accompany speciation in fishes.

5. Summary

A combined analysis of mitochondrial and nuclear sequence data revealed that the genus *Opsaridium* is not a monophyletic clade, but instead includes representatives of the genus *Raiamas* from the broader Congo catchment, including Lake Rukwa and Lake Tanganyika. The results also suggest that the three *Opsaridium* species in Lake Malawi catchment arrived from multiple independent colonisation events, as opposed to radiating within the catchment. Much of the diversity within the *Opsaridium* group is found within the Congo region that appears to have been an important source for diversity in the Zambezi system, including Lake

Malawi. Further resolution of the phylogeny of the African chedrin clade will help to improve our understanding of the role of Congo system in acting as a source of diversity within East and Southern Africa.

Acknowledgements

The work was supported by a Commonwealth Scholarship MWCS-2012-249 to HS, and Royal Society-Leverhulme Trust Africa Awards AA100023 and AA130107 to MG and BPN. We thank A. Smith, G. Turner, J. Swanstrom, C. Katongo, S. Mzighani and TAFIRI researchers for assistance with sample collection. We thank G. Orti and two reviewers for helpful comments.

References

Agnèse, J.F., Teugels, G.G. 2001. The *Bathyclarias–Clarias* species flock. A new model to understand rapid speciation in African great lakes. Cr. Acad. Sci. III-Vie. 324, 683-688.

Ahnelt, H., Herdina, A.N., Metscher, B. D. 2015. Unusual pharyngeal dentition in the African Chedrin fishes (Teleostei: Cyprinindae): Significance for phylogeny and character evolution. Zool. Anz. 255, 85-102.

Banyankimbona, G., Vreven, E., Snoeks, J. 2013. A revision of the genus *Astatoreochromis* (Teleostei, Cichlidae), East-Africa. Eur. J. Taxon., 39, 1-21.

Beadle, L.C. 1974. The inland waters of tropical Africa. An introduction to tropical limnology. Longman Group. London.

Britz, R., Conway, K.W., Rüber, L. 2009. Spectacular morphological novelty in a miniature cyprinid fish, *Danionella dracula* n. sp. Proc. Roy. Soc. B 276, 2179-2186.

Brown, K.J., Rüber, L., Bills, R., Day, J.J. 2010. Mastacembelid eels support Lake Tanganyika as an evolutionary hotspot of diversification. BMC Evol. Biol. 10, 188.

Burke, K., and Gunnell, Y. 2008. The African erosion surface: a continental-scale synthesis of geomorphology, tectonics, and environmental change over the past 180 million years. Geol. Soc. Am. Mem., 201, 1-66.

Chen, W.J., Mayden, R.L. 2009. Molecular systematics of the Cyprinoidea (Teleostei: Cypriniformes), the world's largest clade of freshwater fishes: further evidence from six nuclear genes. Mol. Phylogenet. Evol. 52, 544-549.

447 Cohen, A.S., Stone, J.R., Beuning, K.R., Park, L.E., Reinthal, P.N., Dettman, D., Scholz C.A.,
448 Johnson, T.C. King, J.W., Talbot, M.R., Brown, E.T., Ivory, S.J. 2007. Ecological
449 consequences of early Late Pleistocene megadroughts in tropical Africa. Proc. Natl.
450 Acad. Sci. USA 104, 16422-16427.
451 Cohen, A.S., Van Bocxlaer, B., Todd, J.A., McGlue, M., Michel, E., Nkotagu, H.H., Grove,
452 A.T., Delvaux, D. 2013. Quaternary ostracodes and molluscs from the Rukwa Basin
453 (Tanzania) and their evolutionary and paleobiogeographic implications. Palaeogeogr.
454 Palaeocl. 392, 79-97.
455 Conway, K.W., Chen, W.J., Mayden, R.L. 2008. The "Celestial Pearl danio" is a miniature
456 *Danio* (ss) (Ostariophysi: Cyprinidae): evidence from morphology and
457 molecules. Zootaxa 1686, 1-28.
458 Dalibard, M., Popescu, S.M., Maley, J., Baudin, F., Melinte-Dobrinescu, M.C., Pittet, B.,
459 Marsset, T., Dennielou, B., Droz, L., Suc, J.P. 2014. High-resolution vegetation history
460 of West Africa during the last 145 ka. Geobios 47, 183-198.
461 Danley, P.D., Husemann, M., Ding, B., DiPietro, L.M., Beverly, E.J., Peppe, D.J. 2012. The
462 impact of the geologic history and paleoclimate on the diversification of East African
463 cichlids. Int. J. Evol. Biol. 2012, Article ID 574851.
464 Day, J.J., Peart, C.R., Brown, K.J., Friel, J.P., Bills, R., Moritz, T. 2013. Continental
465 diversification of an African catfish radiation (Mochokidae: *Synodontis*). Syst. Biol.
466 62, 351-365.
467 De Vos, L., Seegers, L. 1998. Seven new *Orthochromis* species (Teleostei: Cichlidae) from the
468 Malagarasi, Luiche and Rugufu basins (Lake Tanganyika drainage), with notes on their
469 reproductive biology. Ichthyol. Explor. Freshwaters 9, 371-420.
470 Dorn, A., Ng'oma, E., Janko, K., Reichwald, K., Polačik, M., Platzer, M., Cellerino, A.,
471 Reichard, M. 2011. Phylogeny, genetic variability and colour polymorphism of an
472 emerging animal model: the short-lived annual *Nothobranchius* fishes from southern
473 Mozambique. Mol. Phylogenet. Evol. 61, 739-749.
474 Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A. 2012. Bayesian phylogenetics with
475 BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969-1973.
476 Egger, B., Kläfiger, Y., Indermaur, A., Koblmüller, S., Theis, A., Egger, S., Näf, T., Van
477 Steenberge, M., Sturmbauer, C., Katongo, C., Salzburger, W. 2015. Phylogeographic and
478 phenotypic assessment of a basal haplochromine cichlid fish from Lake Chila, Zambia.
479 Hydrobiologia 748, 171-184.

480 Eschmeyer, W.N., Fricke, R., van der Laan R. 2016. Catalog of fishes: Genera, species,
481 references. Electronic version accessed 20 June 2016.
482 http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp.
483 Fang, F., Noren, M., Liao, T.Y., Källersjö, M., Kullander, S.O. 2009. Molecular phylogenetic
484 interrelationships of the south Asian cyprinid genera *Danio*, *Devario*,
485 and *Microrasbora* (Teleostei, Cyprinidae, Danioninae). Zool. Scr. 38, 237–256.
486 Genner, M.J., Ngatunga, B.P., Mzighani, S., Smith, A., Turner, G.F. 2015. Geographical
487 ancestry of Lake Malawi's cichlid fish diversity. Biol. Lett. 11, 20150232.
488 Goodier, S.A., Cotterill, F.P., O'Ryan, C., Skelton, P.H., de Wit, M.J. 2011. Cryptic diversity
489 of African tigerfish (Genus *Hydrocynus*) reveals palaeogeographic signatures of linked
490 Neogene geotectonic events. PloS ONE 6, e28775.
491 He, S., Mayden, R.L., Wang, X., Wang, W., Tang, K.L., Chen, W.J., Chen, Y. 2008. Molecular
492 phylogenetics of the family Cyprinidae (Actinopterygii: Cypriniformes) as evidenced by
493 sequence variation in the first intron of S7 ribosomal protein-coding gene: Further
494 evidence from a nuclear gene of the systematic chaos in the family. Mol. Phylogenet.
495 Evol. 46, 818-829.
496 Howes, G.J. 1980. The anatomy, phylogeny and classification of bariliine cyprinid fishes. Bull.
497 British Mus. Nat. Hist. Zool. 37, 129-198.
498 Katongo, C., Koblmüller, S., Duftner, N., Mumba, L., Sturmbauer, C. 2007. Evolutionary
499 history and biogeographic affinities of the serranochromine cichlids in Zambian rivers.
500 Mol. Phylogenet. Evol. 45, 326-338.
501 Koblmüller, S., Katongo, C., Phiri, H, Sturmbauer, C. 2012. Past connection of the upper
502 reaches of a Lake Tanganyika tributary with the upper Congo drainage suggested by
503 genetic data of riverine cichlid fishes. Afr. Zool. 47, 182-186.
504 Koblmüller, S., Sefc. K.M., Sturmbauer, C. 2008a. The Lake Tanganyika cichlid species
505 assemblage: recent advances in molecular phylogenetics. Hydrobiologia 615, 5-20.
506 Koblmüller, S., Schliewen, U.K., Duftner, N., Sefc, K.M., Katongo, C., Sturmbauer, C. 2008b.
507 Age and spread of the haplochromine cichlid fishes in Africa. Mol. Phylogenet. Evol. 49,
508 153-169.
509 Lanfear, R., Calcott, B., Ho, S. Y., Guindon, S. 2012. PartitionFinder: combined selection of
510 partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol.
511 29, 1695-1701.
512 Lévêque, C., Oberdorff, T., Paugy, D., Stiassny, M.L.J., Tedesco, P.A. 2008. Global diversity
513 of fish (Pisces) in freshwater. Hydrobiologia 595, 545–567.

514 Liao, T.Y., Arroyave, J., Stiassny, M.L. 2012. Diagnosis of Asian *Raiamas* (Teleostei:
515 Cyprinidae: Chedrina) with comments on chedrin relationships and previously proposed
516 diagnostic characters for *Opsaridium* and *Raiamas*. Ichthyol. Res. 59, 328–341.
517 Liao, T.-Y., Kullander, S.O., Fang, F. 2011. Phylogenetic position of rasborin cyprinids and
518 monophyly of major lineages among the Danioninae, based on morphological characters
519 (Cypriniformes: Cyprinidae). J. Zool. Syst. Evol. Res. 49, 224–232.
520 Mayden, R.L., Chen, W.J., Bart, H.L., Doosey, M.H., Simons, A.M., Tang, K.L., Wood, R.M.,
521 Agnew, M.K., Yang, L., Hirt, M.V., Clements, M.D., Saitoh, K., Sado, T., Miya, M.,
522 Nishida, M. 2009. Reconstructing the phylogenetic relationships of the earth's most
523 diverse clade of freshwater fishes – order Cypriniformes (Actinopterygii: Ostariophysi):
524 a case study using multiple nuclear loci and the mitochondrial genome. Mol. Phylogenet.
525 Evol. 51, 500-514.
526 Mayden, R.L., Tang, K.L., Wood, R.M., Chen, W.-J., Agnew, M.K., Conway, K.W., Yang, L.,
527 Simons, A.M., Bart, H.L., Harris, P.M., Li, J., Wang, X., Saitoh, K., He, S., Liu, H.,
528 Chen, Y., Nishida, M., Miya, M. 2008. Inferring the Tree of Life of the order
529 Cypriniformes, the earth's most diverse clade of freshwater fishes: Implications of varied
530 taxon and character sampling. J. Syst. Evol. 46, 424-438.
531 Meyer, B.S., Indermaur, A., Ehrensperger, X., Egger, B., Banyankimbona, G., Snoeks, J.,
532 Salzburger, W. 2015. Back to Tanganyika: a case of recent trans-species-flock dispersal
533 in East African haplochromine cichlid fishes. R. Soc. Open Sci. 2, 140498.
534 Moore, A.E., Eckardt, F. 2012. The evolution and ages of Makgadikgadi palaeo-lakes:
535 Consilient evidence from Kalahari drainage evolution south-central Africa. S. Afr. J.
536 Geol. 115, 385-413.
537 Near, T.J., Eytan, R.I., Dornburg, A., Kuhn, K.L., Moore, J.A., Davis, M.P., Wainwright, P.C.,
538 Friedman, M., Smith, W.L. 2012. Resolution of ray-finned fish phylogeny and timing of
539 diversification. Proc. Natl. Acad. Sci. USA 109, 13698-13703.
540 Nichols, P., Genner, M.J., van Oosterhout, C., Smith, A., Parsons, P., Sungani, H., Swanstrom,
541 J., Joyce, D.A. 2015. Secondary contact seeds phenotypic novelty in cichlid fishes. Proc.
542 Roy. Soc. B 282, 20142272.
543 Oberdorff, T., Tedesco, P. A., Hugueny, B., Leprieur, F., Beauchard, O., Brosse, S., Dürr, H.H.
544 2011. Global and regional patterns in riverine fish species richness: a review. Int. J. Ecol.
545 2011, Article ID 967631.

546 Peart, C.R., Bills, R., Wilkinson, M., Day, J.J. 2014. Nocturnal claroteine catfishes reveal dual
547 colonisation but a single radiation in Lake Tanganyika. Mol. Phylogenet. Evol. 73, 119-
548 128.
549 Pinton, A., Agnèse, J.F., Paugy, D., Otero, O. 2013. A large-scale phylogeny of *Synodontis*
550 (Mochokidae, Siluriformes) reveals the influence of geological events on continental
551 diversity during the Cenozoic. Mol. Phylogenet. Evol. 66, 1027-1040.
552 Rambaut, A., Suchard, M., Drummond, A.J. 2013. Tracer v.1.6. <
553 http://tree.bio.ed.ac.uk/software/tracer/>.
554 Renner, S.S., Zhang, L.-B. 2004. Biogeography of the *Pistia* clade (Araceae): based on
555 chloroplast and mitochondrial DNA sequences and Bayesian divergence time inference.
556 Syst. Biol. 53, 422-432.
557 Roberts, T.R. 1975. Geographical distribution of African freshwater fishes. Zool J. Linn Soc.
558 57, 249–319.
559 Robinson, J., Schmitt, E.A., Dowling, J. E. 1995. Temporal and spatial patterns of opsin gene
560 expression in zebrafish (*Danio rerio*). Vis. Neurosci. 12, 895-906.
561 Rüber, L., Kottelat, M., Tan, H.H., Ng, P.K., Britz, R. 2007. Evolution of miniaturization and
562 the phylogenetic position of *Paedocypris,* comprising the world's smallest
563 vertebrate. BMC Evol. Biol. 7, 1.
564 Saitoh, K., Sado, T., Mayden, R.L., Hanzawa, N., Nakamura, K., Nishida, M., Miya, M. 2006.
565 Mitogenomic evolution and interrelationships of the Cypriniformes (Actinopterygii:
566 Ostariophysi): The first evidence toward resolution of higher-level relationships of the
567 world's largest freshwater fish clade based on 59 whole mitogenome sequences. J. Mol.
568 Evol. 63, 826-841.
569 Salzburger, W., Van Bocxlaer, B., Cohen, A.S. 2014. Ecology and evolution of the African
570 Great Lakes and their faunas. Ann. Rev. Ecol. Evol. Syst. 45, 519-545.
571 Schmidt, R.C., Bart, H.L., Nyingi, D.W., Gichuki, N.N. 2014. Phylogeny of suckermouth
572 catfishes (Mochokidae: *Chiloglanis*) from Kenya: The utility of Growth Hormone introns
573 in species level phylogenies. Mol. Phylogenet. Evol. 79, 415-421.
574 Seehausen, O. 2015. Process and pattern in cichlid radiations–inferences for understanding
575 unusually high rates of evolutionary diversification. New Phytologist 207, 304–312.
576 Shimodaira H., Hasegawa, M. 2001. CONSEL: for assessing the confidence of phylogenetic
577 tree selection. Bioinformatics 17, 1246-1247.
578 Skelton, P.H. 1996. A review of *Opsaridium zambezense* from southern Africa with a
579 description of a new species from Malawi. Ichthyol. Explor. Freshwaters 7, 59-84.

580 Skelton, P.H. 2001. A complete guide to the freshwater fishes of southern Africa. Struik, Cape
581 Town.
582 Šlechtová, V., Bohlen, J., & Tan, H.H. (2007). Families of Cobitoidea (Teleostei;
583 Cypriniformes) as revealed from nuclear genetic data and the position of the mysterious
584 genera *Barbucca*, *Psilorhynchus*, *Serpenticobitis* and *Vaillantella*. Mol. Phylogenet.
585 Evol. 44, 1358-1365.
586 Snoeks J. 2004. The non-cichlid fishes of the Lake Malawi/Nyasa/ Niassa system, in: Snoeks
587 J. (Ed.), The cichlid diversity of Lake Malawi/Nyasa/Niassa: identification, distribution
588 and taxonomy. Cichlid Press, El Paso, Texas, pp. 20–26.
589 Snoeks, J., Harrison, I.J., Stiassny, M.L.J. 2011. The status and distribution of freshwater
590 fishes, in Darwall, W.R.T., Smith, K.G., Allen, D.J., Holland, R.A., Harrison, I.J.,
591 Brooks, E.G.E. (eds), The diversity of life in African freshwaters: Under Water, Under
592 Threat. An analysis of the status and distribution of freshwater species throughout
593 mainland Africa. IUCN, Cambridge UK/Gland Switzerland, pp. 42-73.
594 Stamatakis, A., Hoover, P., Rougemont, J. 2008. A rapid bootstrap algorithm for the RAxML
595 web servers. Syst. Biol. 57, 758-771.
596 Stankiewicz J., de Wit M.J. 2006. A proposed drainage evolution model for Central Africa—
597 Did the Congo flow east?. J Afr. Earth. Sci. 44, 75-84.
598 Stewart, K.M., Murray, A.M. 2013. Earliest fish remains from the Lake Malawi Basin, Malawi,
599 and biogeographical implications. J. Vert. Paleont. 33, 532-539.
600 Stiassny, M.L.J., Schelly R.C., Schliewen. U.K. 2006. A new species of *Raiamas* (Teleostei:
601 Cyprinidae) from the Lower Congo River, with a phylogenetic assessment of the generic
602 limits of the predatory cyprinid genera *Opsaridium, Raiamas* and *Leptocypris*. Copeia
603 2006, 370-377.
604 Tang, K.L., Agnew, M.K.M., Hirt, V., Sado, T., Schneider, L.M., Freyhof, J., Sulaiman, Z.,
605 Swartz, E., Vidthayanon, C., Miya, M., Saitoh, K., Simons, A.M., Wood, R.M., Mayden,
606 R.L. 2010. Systematics of the subfamily Danioninae (Teleostei: Cypriniformes:
607 Cyprinidae). Mol. Phylogenet. Evol. 57, 189-214.
608 Tweddle, D., Lewis, D.S.C., Willoughby, N.G. 1979. The nature of the barrier separating the
609 Lake Malawi and Zambezi fish faunas. Ichthyol. Bull. Rhodes Univ. 39, 1–9
610 Tweddle, D., Skelton, P.H. 2008. New species of '*Barbus*' and *Labeobarbus* (Teleostei:
611 Cyprinidae) from the South Rukuru River, Malawi, Africa. Smithiana Bull. 8, 25-49.
612 Tweddle, D., Turner, G.F. 2014. Evidence for lacustrine breeding by sanjika *Opsaridium*
613 *microcephalum* (Teleostei: Cyprinidae) in Lake Malawi. Afr. J. Aquat. Sci. 39, 479-480.

614 Tweddle, D., Willoughby, N.G. 1979. An annotated checklist of the fish fauna of the River
615 Shire south of Kapachira Falls, Malawi. Ichthyol. Bull. Rhodes Univ. 39, 11–22.
616 Willett, C.E., Zapata, A.G., Hopkins, N., Steiner, L.A. 1997. Expression of zebrafish rag genes
617 during early development identifies the thymus. Dev. Biol. 182, 331-341.
618 Xia, X. 2013. DAMBE5: A comprehensive software package for data analysis in molecular
619 biology and evolution. Mol. Biol. Evol. 30, 1720-1728.
620 Yu, Y., Harris, A.J., Blair, C., He, X. 2015. RASP (Reconstruct Ancestral State in
621 Phylogenies): a tool for historical biogeography. Mol. Phylogenet. Evol. 87, 46-49.
622

623 **Figure Legends**

624

625 **Fig. 1.** Time calibrated phylogeny of chedrins inferred from Bayesian analysis of mitochondrial
626 and nuclear DNA sequences. Blue node bars represent the 95% highest posterior density (HPD)
627 intervals of inferred divergence time estimates. Values at nodes indicate Bayesian posterior
628 probabilities and maximum likelihood bootstrap percentages, respectively. Underlined taxa are
629 from the Lake Malawi catchment.

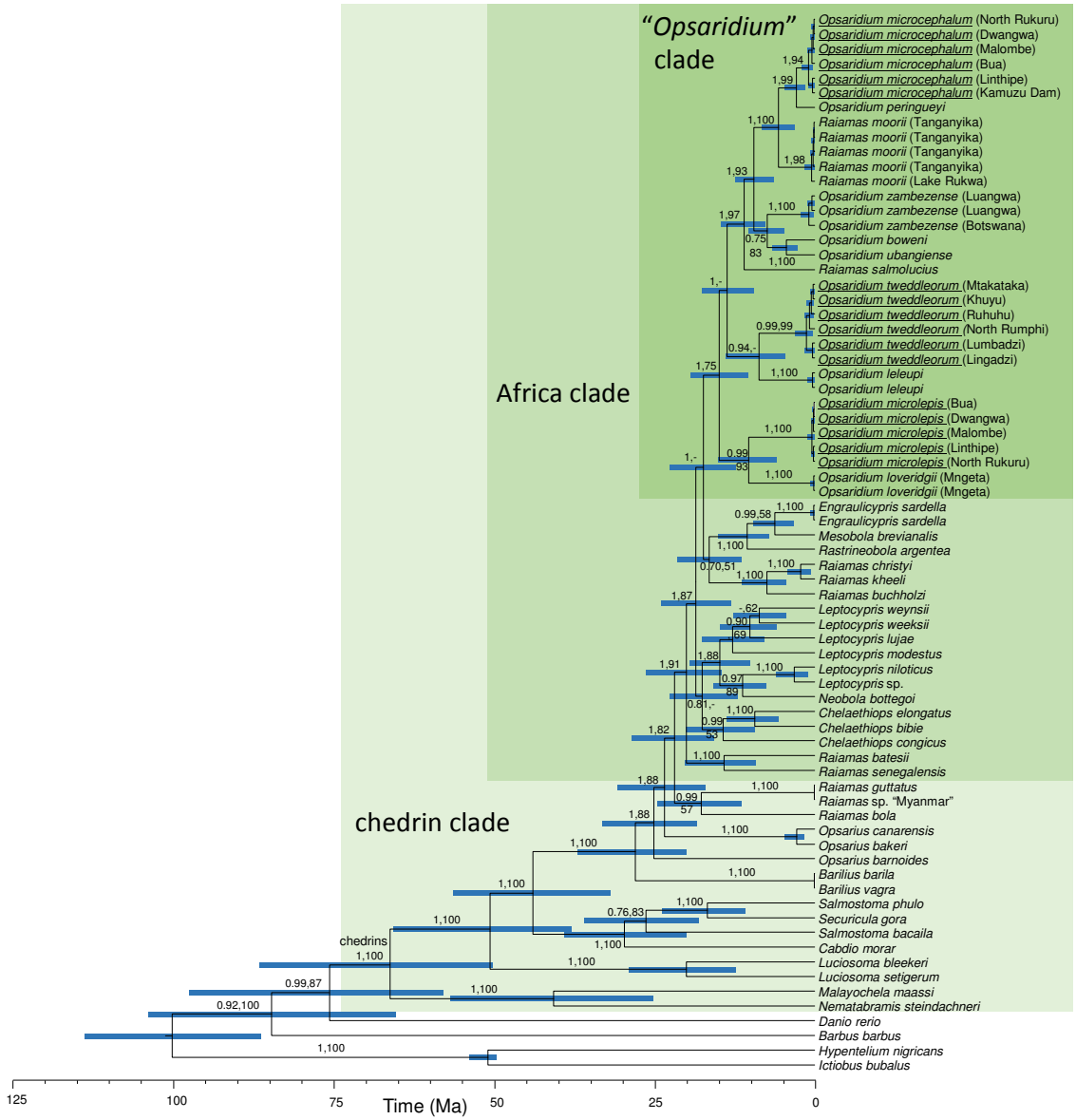
630

631 **Fig. 2.** Distribution of species within the "*Opsaridium*" clade across Africa. Data sourced from
632 GBIF (http://www.gbif.org/), downloaded 31 July 2015. Records conflicting with
633 distributional information of *O. zambezense* in Skelton (1996) and Skelton (2001) were
634 removed.

635

636 **Fig. 3**. Ancestral area reconstruction using the Bayes Lagrange Statistical Dispersal-
637 Extinction-Cladogenesis (S-DEC) analyses as implemented in RASP (Yu *et al.* 2015).
638 Ancestral distribution range with a low percentage likelihood (<15%) are merged for each node
639 and coloured black. Coloured circles indicate the percentage probability of area occupancy.
640 Multiple colours in a segment indicate a distribution shared between areas.

641

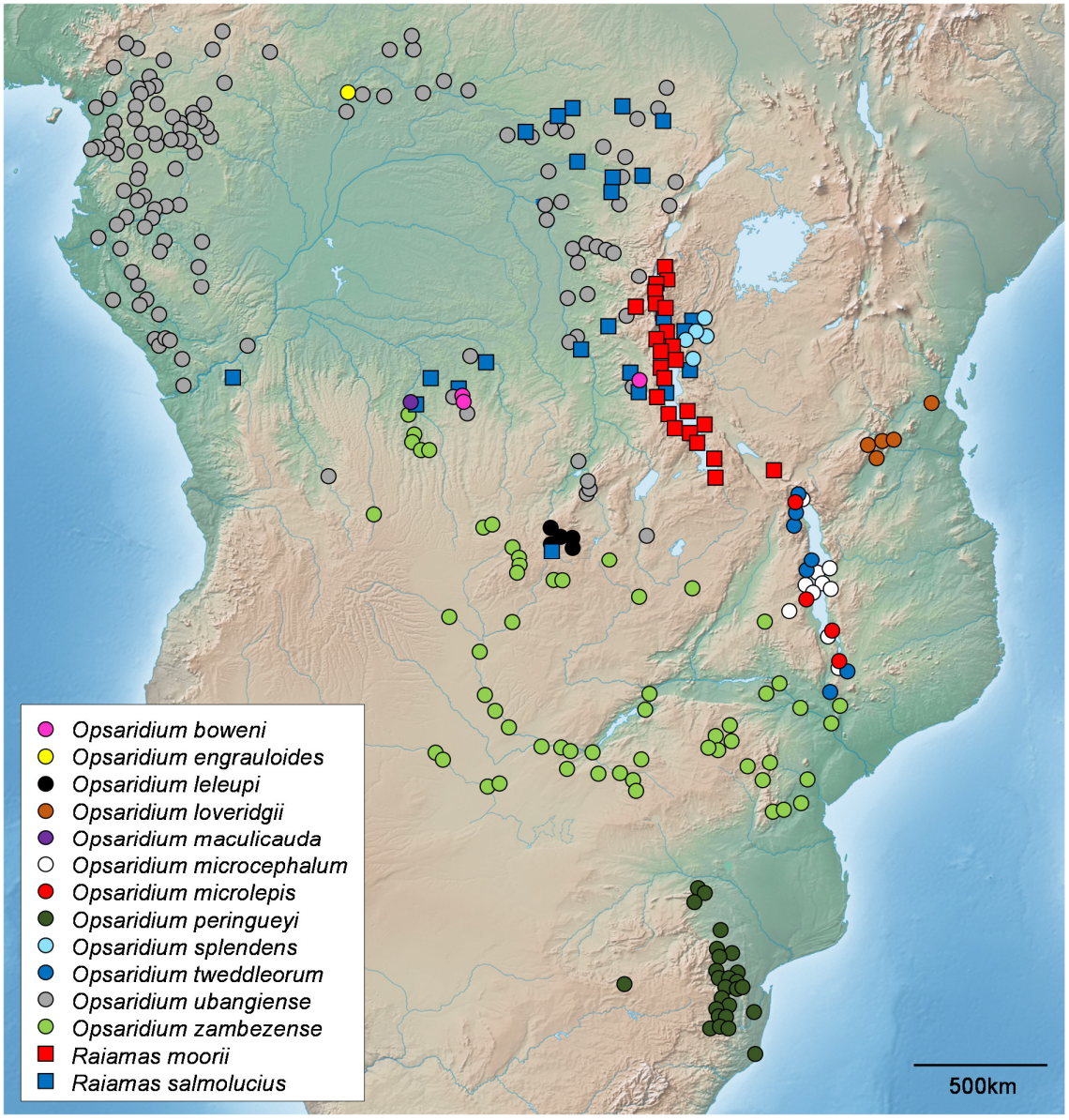


642

643

644 Fig. 1

645

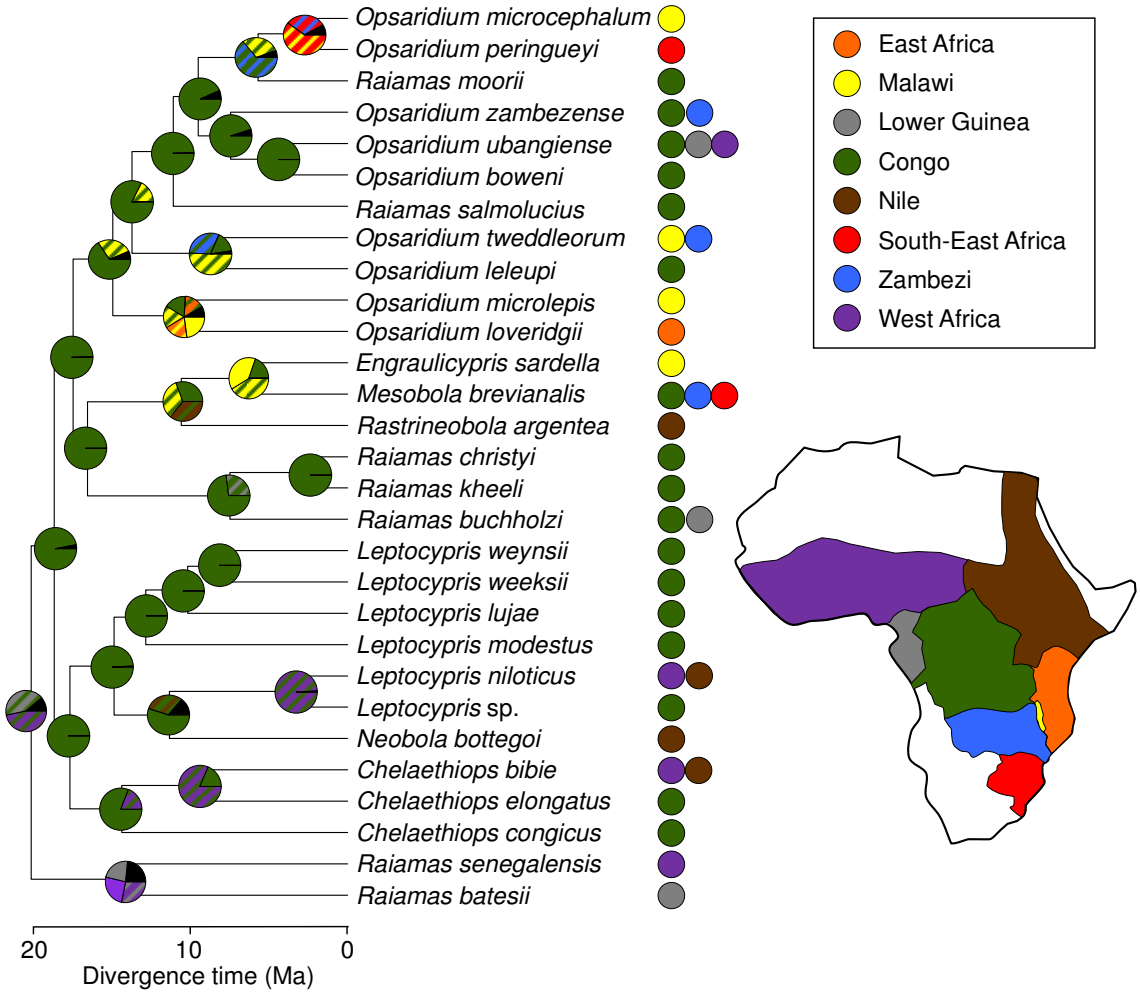


646

647

648 Fig. 2

649



650

651

652 Fig. 3

653 **Table 1**: Collection information (coordinates in degrees and decimal minutes) and Genbank accessions for specimens newly sequenced for this
654 study. n/a indicates no sequence was available.

Species	Location	Catchment	Latitude (S)	Longitude (E)	Date	*COI* Accession	*RAG1* Accession
Opsaridium leleupi	Kisanfu River (a)	Congo	10º45.840'	25º57.840'	09-09-2008*	KX810223	n/a
	Kisanfu River (b) Kicrossing	Congo	10º47.970'	25º58.859'	09-12-2008*	KX810224	n/a
Opsaridium loveridgii	Mngeta (a)	Rufiji (Kilombero)	08º20.457'	36º05.227'	07-03-2014	KX810212	KX810226
	Mngeta (b)	Rufiji (Kilombero)	08º20.457'	36º05.227'	07-03-2014	KX810213	n/a
Opsaridium microcephalum	Malombe	Lake Malawi	14º32.580'	35º12.660'	04-05-2010	KX810218	n/a
	Bua	Lake Malawi	12º47.194	34º11.755'	24-01-2014	KU941646	n/a
	Dwangwa	Lake Malawi	12º30.807'	34º06.971'	25-01-2014	KX810216	n/a
	Linthipe	Lake Malawi	13º47.333'	34º26.088'	22-01-2014	KU941590	n/a
	Kamuzu Dam	Lake Malawi	14º10.465'	33º38.566'	23-08-2013	KX810217	n/a
	North Rukuru	Lake Malawi	09º55.032'	33º55.676'	28-01-2014	KU941673	KX810228
Opsaridium microlepis	Malombe	Lake Malawi	14º32.580'	35º12.660'	04-05-2010	KU941774	n/a
	Dwangwa	Lake Malawi	12º30.807'	34º06.971'	25-01-2014	KU941767	n/a
	Bua	Lake Malawi	12º47.194'	34º11.755'	24-01-2014	KU941712	n/a
	North Rukuru	Lake Malawi	09º55.032'	33º55.676'	28-01-2014	KU941758	KX810230
	Linthipe	Lake Malawi	13º47.333'	34º26.088'	22-01-2014	KU941691	KX810229
Opsaridium tweddleorum	Mtakataka	Lake Malawi	14º12.855'	34º30.819'	12-01-2011	KU941820	n/a
	Khuyu	Lake Malawi	12º16.735'	33º59.939'	26-01-2014	KU941806	n/a
	Ruhuhu	Lake Malawi	10º37.603'	35º39.192'	09-09-2012	KU941831	n/a
	North Rumphi	Lake Malawi	10º41.109'	34º10.883'	27-01-2014	KU941793	KX810236
	Lumbadzi	Lake Malawi	13º47.640'	33º59.300'	23-08-2013	n/a	KX810235
	Lingadzi	Lake Malawi	13º32.061'	34º17.064'	23-01-2014	n/a	KX810234
Opsaridium zambezense	Lutembe (a)	Zambezi (Luangwa)	13º42.383'	32º29.366'	21-06-2010	KX810214	KX810227
	Lutembe (b)	Zambezi (Luangwa)	13º42.383'	32º29.366'	21-06-2010	KX810215	n/a
Raiamas moorii	Kavimira River (a)	Lake Tanganyika	03º21.156'	29º09.054'	10-09-2013	KX810221	n/a
	Kavimira River (b)	Lake Tanganyika	03º21.156'	29º09.054'	10-09-2013	KX810220	n/a
	Kalimabenge River (a)	Lake Tanganyika	03º24.434'	29º07.585'	10-09-2013	KX810222	KX810233
	Kalimabenge River (b)	Lake Tanganyika	03º24.434'	29º07.585'	10-09-2013	KX810219	KX810231
	Lake Rukwa	Lake Rukwa	08º 23.850	32º54.110'	01-09-2012	n/a	KX810232
Engraulicypris sardella	Lake Malawi	Lake Malawi	14º04.172'	34º55.034'	10-01-2015	KX810225	TBC

655 *collectors: K. Magellan, J. Harvey, G. Winnaar

656 **Table 2**: Taxa, voucher specimens, and accession numbers for the gene sequences from GenBank used in this study. n/a indicates no sequence
657 was available.

Taxon group	Species	*COI*	*cytb*	*RAG1*	*RH1*	References
African chedrins	*Chelaethiops bibie*	HM223902	HM224023	HM224260	JX196995	Tang et al. 2010, Liao et al. 2012
	Chelaethiops cf. congicus	HM418170	n/a	n/a	n/a	International Barcode of Life submission
	Chelaethiops congicus	HM880211	n/a	n/a	n/a	International Barcode of Life submission
	Chelaethiops elongatus	JX196996	JX197006	JX197014	JX197021	Liao et al. 2012
	Chelaethiops elongatus	JX196996	JX197021	JX197014	JX197006	Liao et al. 2012
	Engraulicypris sardella	JX196997	JX197022	JX197015	JX197007	Liao et al. 2012
	Leptocypris lujae	KT193306	n/a	n/a	n/a	Decru et al. unpub
	Leptocypris modestus	JX196998	JX197008	HM224052	HM223933	Tang et al. 2010, Liao et al. 2012
	Leptocypris niloticus	JX196999	HM224293	HM224050	HM223932	Tang et al. 2010, Liao et al. 2012
	Leptocypris sp.	n/a	n/a	HM224052	HM223933	Tang et al. 2012
	Leptocypris weeksii	HM418175	n/a	n/a	n/a	International Barcode of Life submission
	Leptocypris weynsii	HM418188	n/a	n/a	n/a	International Barcode of Life submission
	Mesobola brevianalis	HM224176	HM224295	HM224054	HM223935	Tang et al. 2010
	Neobola bottegoi	HM224178	HM224296	HM224056	HM223936	Tang et al. 2010
	Opsaridium boweni	JX197000	JX197009	JX197016	JX197023	Liao et al. 2012
	Opsaridium peringueyi	HM224192	HM224311	HM224072	HM223954	Tang et al. 2010
	Opsaridium ubanguiense	JX197001	HM224312	HM224073	HM223956	Tang et al. 2010, Liao et al. 2012
	Opsaridium zambezense	HM224194	n/a	n/a	n/a	Tang et al. 2010
	Raiamas batesii	JX197002	JX197010	JX197017	JX197024	Liao et al. 2012
	Raiamas buchholzi	HM224213	HM224330	HM224090	HM223975	Tang et al. 2010
	Raiamas cf. christyi	HM418152	n/a	n/a	n/a	International Barcode of Life submission
	Raiamas cf. christyi	KT193013	n/a	n/a	n/a	Decru et al. unpublished
	Raiamas christyi	JX197003	JX197011	JX197018	JX197025	Liao et al. 2012
	Raiamas kheeli	HM418197	n/a	n/a	n/a	International Barcode of Life submission
	Raiamas salmolucius	JX197004	JX197012	JX197019	JX197026	Liao et al. 2012
	Raiamas senegalensis	HM224215	HM224332	HM224093	HM223978	Tang et al. 2010
	Rastrineobola argentea	JX197005	n/a	JX197020	JX197027	Liao et al. 2012
Asiatic chedrins	*Barilius barila*	HM224138	HM224257	HM224020	HM223900	Tang et al. 2010
	Barilius vagra	HM224140	HM224259	HM224022	HM223901	Tang et al. 2010
	Cabdio morar	AP011335	AP011335	FJ531343	EU711105	Tang et al. 2010, Chen & Mayden 2009
	Luciosoma bleekeri	AP011399	AP011399	HM224053	HM223934	Tang et al. 2010, Liao et al. 2012
	Luciosoma setigerum	AP011423	AP011423	EU292704	FJ531352	Conway et al. 2008, Chen & Mayden 2009, Tang et al. 2010.

	Malayochela maassi	FJ753486	EF151098	FJ753522	EU241379	Rüber et al. 2007, Britz et al. 2009, Fang et al. 2009
	Nematabramis steindachneri	FJ753496	EF151106	FJ753532	n/a	Rüber et al. 2007, Britz et al. 2009,
	Opsarius bakeri	HM224197	HM224315	HM224076	HM223959	Tang et al. 2010
	Opsarius barnoides	HM224199	HM224317	HM224078	HM223961	Tang et al. 2010
	Opsarius canarensis	HM224200	HM224318	HM224079	HM223962	Tang et al. 2010
	Raiamas bola	HM224212	HM224329	HM224089	HM223974	Tang et al. 2010
	Raiamas guttatus	HM224214	HM224331	HM224091	HM223976	Tang et al. 2010
	Raiamas sp. ‘‘Myanmar’’	HM224217	HM224334	HM224095	HM223980	Tang et al. 2010
	Salmostoma bacaila	AP011223	AP011223	HM224128	HM224010	Tang et al. 2010
	Salmostoma phulo	HM224248	HM224379	HM224133	HM224013	Tang et al. 2010
	Securicula gora	HM224250	HM224381	HM224135	HM224015	Tang et al. 2010
Outgroups	*Barbus barbus*	AB238965	AB238965	EU711147	FJ197049	Saitoh et al. 2006, Mayden et al. 2008, Mayden et al. 2009
	Danio rerio	AC024175	AC024175	U71093	AF105152	Milam et al. unpublished, Willett et al. 2007, Robinson et al. 1995
	Hypentelium nigricans	KF929987	n/a	EU711134	n/a	Bentley and Wiley unpublished, Mayden et al. 2008
	Ictiobus bubalus	KF929996	n/a	EF056353	n/a	Bentley and Wiley unpublished, , Slechtova et al. 2007

658
659